Volume 267, number 1, 1–5 FEBS 08539 July 1990

Structural analysis of the regulatory region of the human corticotropin releasing hormone gene

Nikos C. Vamvakopoulos[1,2], Michael Karl[1], Vanessa Mayol[1], Teresa Gomez[1], Constantine A. Stratakis[1], Andrew Margioris[1] and George P. Chrousos[1]

[1] *Pediatric Endocrinology Section, National Institute of Child Health and Human Development, Bethesda, MD, USA* and [2] *Eucaryotic Gene Expression Laboratory, Institute of Biology, Democritos, Athens, Greece*

Received 18 March 1990; revised version received 21 April 1990

A DNA fragment containing the human corticotropin releasing hormone (CRH) gene, along with 9 kb of upstream and 4 kb of downstream sequences, was isolated from a human genomic DNA library. Nucleotide sequence analysis of the proximal 918 nucleotides 5′ flanking the putative major mRNA start site of the human gene and comparison to the 866 nucleotide long homologous ovine sequence, revealed that this region of the CRH gene consists of two distinct areas with different degrees of homology, varying from 72% to 94%. The putative functional features of the human sequence were identified. Many, but not all, features were conserved in the ovine sequence. The highly conserved nature of the regulatory region of this gene makes it a good candidate for tracing possible related genetic defects of the hypothalamic-pituitary-adrenal (HPA) axis

Corticotropin releasing hormone; Genomic DNA cloning; Nucleotide sequence; Regulatory region

1. INTRODUCTION

Corticotropin releasing hormone (CRH), a 41 amino acid peptide, plays a pivotal role in the maintenance of homeostasis and in the adaptive response to stress [1–3]. Both hypersecretion and hyposecretion of CRH have been described, respectively, in patients with major psychiatric disease, such as depression and anorexia nervosa [4,5], and in a rat model of autoimmune arthritis [6,7]. The CRH gene directs the biosynthesis of hypothalamic, suprahypothalamic and placental pre-proCRH, by complex transient and/or tissue-specific regulatory mechanisms [8–12]. Understanding these mechanisms is a prerequisite to a thorough comprehension of CRH regulation at the basal and stress states and its potential role in the pathophysiology of human disease.

We undertook to delineate the structure of the regulatory region of the human CRH gene as a first step in tracing possible genetic defects of the hypothalamic-pituitary-adrenal (HPA) axis. Previous work revealed that the first 330 based 5′ flanking the human, ovine and rat CRH genes had a higher degree of homology than their protein-coding regions [13,14]. To define the limits of the highly conserved 5′ flanking segment, we isolated a clone from a human genomic DNA library which contains the CRH gene and 9 kb of its upstream and 4 kb of its downstream flanking portions. We sequenced the proximal 918 nucleotides 5′ flanking the putative proximal mRNA start site of the human CRH gene and compared it with that of the ovine gene [14]. The 5′ flanking regions of both genes consist of two distinct areas of sequence homology. The 334 base-long one, proximal to the mRNA start site, with 94% homology and the remaining upstream sequence with 72% homology. The putative functional features of the regulatory region of the human CRH gene are reported along with their extent of preservation at homologous positions of the ovine CRH sequence.

Correspondence address: G.P. Chrousos, NIH, Bld. 10, 10N-262, Bethesda, MD 20892, USA

Abbreviations: CRH, corticotropin-releasing hormone; pre-pro CRH, corticotropin releasing hormone precursor; HPA axis, hypothalamic-pituitary-adrenal axis; POMC, proopiomelanocortin; RFR, regulatory flanking region

2. MATERIALS AND METHODS

2.1. *Genomic library screening and Southern blot analysis*

The human genomic DNA library, constructed in a charon 30 vector [15], has been previously described [16,17]. The probe used for screening was a synthetic 48-mer oligodeoxynucleotide complementary to human CRH mRNA [9]. This oligomer was radiolabeled at its 5′ end with [γ-P^{32}]ATP (3000 Ci/mmol, ICN) and T4 polynucleotide kinase, using standard procedures [18]. The radiolabeled probe was purified by Sephadex G-50 gel filtration and was used throughout the identification, purification and characterization of the human CRH gene-containing genomic clone. The specific activity of the column-purified oligonucleotide probe was 2×10^8 cpm/nmol. All the prehybridizations and hybridizations were done at 42°C in a solution containing 1 M sodium chloride, 50% formamide and 10% dextran sulfate (Pharmacia), pH 7. Only the prehybridization solutions contained 100 μg/ml boiled and sonicated salmon sperm DNA carrier. Prior to autoradiography, the blots were washed twice at 55°C for 30 min to a final stringency of 2 times standard saline citrate/0.1% sodium dodecyl sulfate. The *Bst*EII-digested DNA, used as a size

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

standard, was radiolabeled at the 3′ ends with the Klenow fragment of DNA polymerase using [α-^{32}P]dGTP (800 Ci/mmol, ICN) as the label, as described in [18]

2.2. *DNA sequencing*

The hybridizing 3.8 kb *Hin*dIII fragment containing the human CRH gene [19], was subcloned at both orientations on a bluescript plasmid vector (Stratagene). DNA sequencing was performed using T7 DNA polymerase and dideoxy-termination reactions [20]. It was carried out with a commercial kit (US Biochemicals, kit version 2.0), using [α-^{35}S]dATP (1000 Ci/mmol, Amersham), as the label. Sequencing of both strands was performed using the following CRH-specific primers: 5′GAGATTTCCCAAGTGTG3′, 5′CAAGCACATAAAGGAG3′, and 5′GATGTTCCTGCACAC3′, in addition to the T3 primer.

2.3. *Materials*

Most of the biological reagents were obtained from Bethesda Research Laboratories and Boehringer Mannheim and the solutions were prepared by the NIH media unit using published formulations [18].

3. RESULTS AND DISCUSSION

A genomic DNA library derived from human placental DNA was screened with an oligonucleotide probe complementary to human CRH mRNA corresponding to amino acids 22–37 of the published sequence [19]. From about 1.5×10^6 recombinant plaques screened [21], one hybridization-positive clone was isolated. The isolate had a 15 kb DNA insert containing the 3.8 kb *Hin*dIII fragment that carries the human CRH gene. Specifically, it had 9 kb of upstream and 4 kb of downstream sequences surrounding the gene. Fig. 1 shows the *Hin*dIII map of the clone, as derived by hybridization to the CRH oligomer prober after restriction digestion, agarose gel electrophoresis, and Southern transfer of the resulting fragments. The hybridizing 3.8 kb *Hin*dIII fragment was subcloned in a bluescript plasmid vector and the sequence of the 918 base-long 5′ flanking region of the gene was determined (Fig. 2).

The sequence of the 918 base-long 5′ flanking portion of the human CRH gene has been previously determined up to −334 [19]. The regulatory features of the latter, which include the two highlighted TATAA and three CAAT boxes and cAMP response element, have been previously identified. It appears that both promoters are active in human placenta and in mouse corticotroph tumor cells stably transfected with the human gene, and that the distal upstream promoter is less efficiently utilized than the proximal one [22,23]. The homologous rat cAMP response element had also been shown to be functional in transient and stable expression systems [24,25].

The highlighted polypurine tract at −829, previously identified at the −801 position of ovine CRH [14], is also present in the −400 bp 5′ flanking region of the proopiomelanocortin [26] and rat growth hormone genes [27], as well as in the mouse mammary tumor virus [28]. A functional role for this conserved sequence has yet to be assigned. A long open reading frame in the same direction with the CRH coding strand exists starting from nucleotide −597 and ending at −224. It

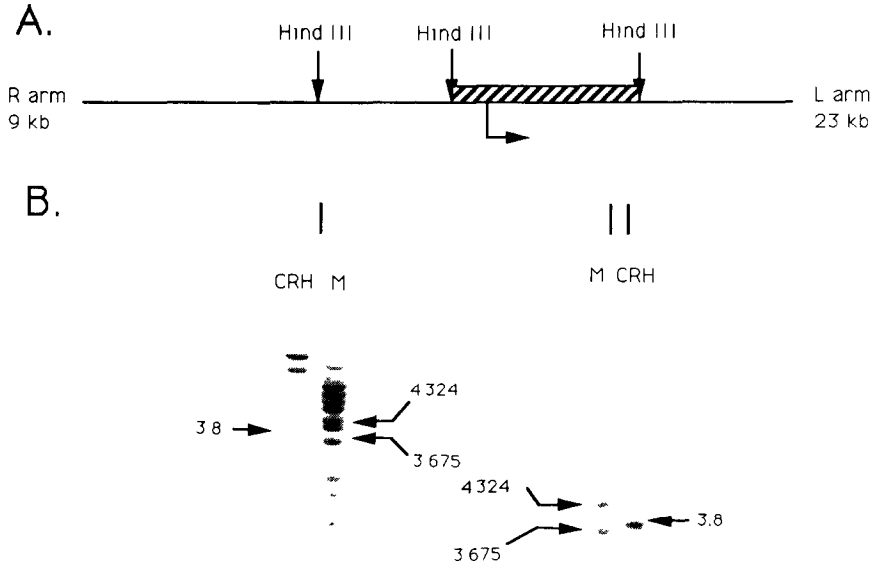

Fig. 1. Positive identification and characterization of the human CRH gene containing genomic DNA isolate. (A) The *Hin*dIII map showing the location of the CRH gene and the extent of its flanking regions in the clone. The bent arrow points to the direction of transcription. The 3.8 kb *Hin*dIII fragment subcloned in the plasmid vector is shown hatched. (B) Southern blot hybridization analysis of the *Hin*dIII-digested hybridization-positive recombinant phage. (I) Picture of ethidium bromide stained gel. (II) Autoradiogram of blotted gel hybridized with the human CRH oligonucleotide probe. Lane M contains the *Bst*EII-digested DNA markers. Lane CRH contains the *Hin*dIII-digested recombinant phage DNA.

Fig. 2. Nucleotide sequence of the complete 5′ flanking region of the human CRH gene within the subcloned 3.8 kb *Hin*dIII fragment. The sequences in bold type letters are the standard regulatory signals of the gene, and the polypurine tract at -829 is a highly conserved feature common to many genes. The sequence shown is colinear to the CRH mRNA coding strand. Underlined is the cyclic AMP-responsive element sequence consensus.

could potentially code for a 125 amino acid-long peptide. Such a long reading frame is not preserved in the ovine gene. Other short open reading frames exist in the human CRH gene which could code for a maximum of 35 amino acids. Comparison of the regulatory region of the human CRH sequence with other sequences from the gene bank revealed a small degree of homology with the proximal (– 300 bp) 5′ regulatory region of several genes, including mink and bovine proopiomelanocortin (POMC). No sequence repetitions or highly repetitive *Alu*I type DNA sequences were identified.

A remarkable structural feature of the CRH gene is the unusually high degree of inter-species sequence homology found in its 5′ flanking region [13,14,25]. To define the limits of the highly conserved 5′ flanking area, we aligned the human and ovine sequences (Fig. 3A). This marked the boundaries of two distinctly conserved regions: regulatory flanking region 1 (RFR 1), with 94% sequence homology, and the adjoining regulatory flanking region 2 (RFR 2), with 72% sequence homology. The degree of homology found in the 5′ RFR 2 of the human CRH gene is similar to that of the 300 bp 5′ flanking regions of other genes, such as the human-mouse and human-bovine POMC [26]. Fig. 3B gives a schematic representation of the two

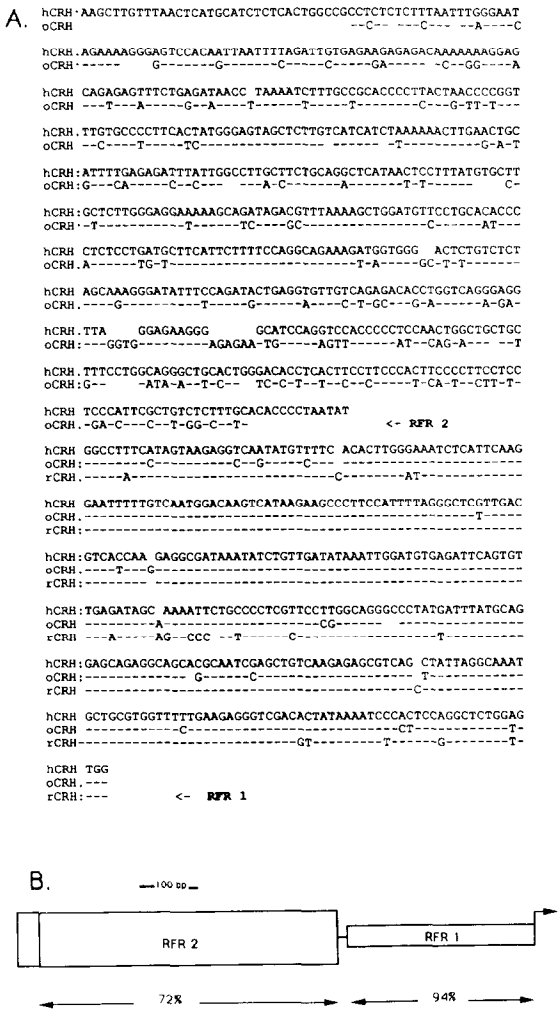

Fig. 3. Alignment of all known CRH regulatory sequences. (A) The human and ovine CRH regulatory sequences were aligned to allow maximal degree of homology. They are displayed in two segments, RFR 1 with 94%, and RFR 2 with 72% overall homologies. The human-rat alignment over the RFR 1 segment is also shown. (B) Schematic representation of human RFR 1 and RFR 2 segments, and their percent homology to ovine segments.

RFR segments linkered by a 13 bp fragment in the human gene that is not present in the ovine one. The high and unusually extended homology of the human CRH 5′ flanking region up to or more than 0.9 kb suggests that the regulation of this gene is complex and influenced by multiple factors. Further upstream sequencing would be important to define the borders of the regulatory region of this gene.

Glucocorticoids have been unequivocally shown to down-regulate rat and ovine hypothalamic CRH mRNA production [14,29,30] and CRH secretion [12]. Interestingly, however, human placental CRH expression was up-regulated by glucocorticoids [9–11]. Although the mechanisms for this phenomenon are still obscure, it is clear that the CRH gene is regulated by glucocorticoids and should contain in its 5′ flanking segment sequences homologous to the steroid receptor binding site consensus [32–34]. In the 5′ RFR 2 of the ovine CRH gene sequence, three such putative signals have been proposed based on their homology to the consensus sequence of one of the two DNA binding zink fingers of the glucocorticoid receptor [33]. These sites, however, have not been preserved in the human sequence, neither is any other new site identifiable with the proposed CRH 'consensus' sequence TGTTCT. In addition, no absolute homology to any other steroid receptor binding consensus sequence was found. Such sites, if present, might reside further upstream and their identification and localization will require further sequencing and appropriate functional assays.

Search for other signal sequences, revealed nine c-jun/AP-1 binding sites, two CP-1 sites, one CTF, and four CACCC sites. Many of these sites are preserved at homologous positions in the ovine CRH sequence. No recognizable sites for Sp1, mammalian enhancer binding protein (C/EBP), AP-2, octamer transcription factor OCT-1 or -2, pituitary factor (pit-1) or serum response factor (SRF) were found [33–35]. Fig. 4, displays a composite of these sites in the upstream region of the human CRH gene. The abundance of c-jun/AP-1 sites, together with their preservation at homologous positions of the ovine gene, is consistent to the experimentally demonstrated inducibility of the CRH gene by phorbol esters and epidermal growth factor [17,36] (and our unpublished results from transient expression studies). It is interesting to note that the CACCC transcription factor appears, under certain conditions, to act synergistically with steroid receptors eliciting strong responses to glucocorticoid stimulation [34]. The CACCC sites found in the human sequence, including the one conserved in the ovine gene, are located in the more heterogeneous RFR 2 domain. It is possible that this binding site might be involved in steroid regulation, if indeed the long sought after steroid response element is located further upstream.

We conclude that the 918 base-long 5′ regulatory region of the human CRH gene consists of two distinct areas with different degrees of evolutionary change. The highly conserved nature of the regulatory region of this gene makes it a good candidate for tracing possibly related genetic defects of the HPA axis. Although consensus sequences for several regulatory factors were recognised, the complexity of CRH regulation in vivo suggests that additional CRH *cis* regulatory elements may exist and for their identification further cloning and functional assays will be required.

A. 5'->

AAGCTTGTTTAACTCATGCATCTCTCACTGGCCGCCTCTCTCTTTAATTTGGGAATAGAAAAGGGAGTCCAC
AATTAATTTTAGATTGTGAGAAGAGAGACAAAAAAAGGAGCAGAGAGTTTCTGAGATAACCTAAAATCTT
TGCCG**CACCC**CTTACTAACCCCGGTTTGTGCCCCTTCACTATGGGAGTAGCTCTTGTCATCATCTAAAAAA
CTTGAACTGCATTTTGAGAGATTTA**TTGG**CCTTGCTTCTGCAGGCTCATAACTCCTTTATGTGCTTGCTCTT
GGGAGGAAAAAGCAGATAGACGTTTAAAAGCTGGATGTTCTGCA**CACCC**CTCTCCTGATGCTTCATTCTTTT
CCAGGCAGAAAGATGGTGGGACTCTGTCTCTAGCAAAGGGATATTTCCAGATACTGAGGTGTT**GTCA**GAGAC
ACCTG**GTCA**GGGAGGTTAGGAGAAGGGGCATCCAGGTC**CACCC**CCTCCAACTGGCTGGCTGCTGCTTTCCTGG
CAGGGCTGCACTGGGACACCTCACTTCTTTCCCACTCCCCCTTCCTCCTCCCATTCGCTGTCTTCTTGCA**CAC**
CCCTAATAT <- **RFR 2**

GGCCTTTCATAGTAAGAG**GTCA**ATATGTTTTCACACTTGGGAAATCTCATTCAAGAATTTTT**GTCA**ATGGA
CAA**GTCA**TAAGAAGCCCTTCCATTTTAGGGCTCGT**TGACGTCA**CCAAGAGGCGATAAATATCTGTTGATA
TAA**ATTGG**ATGTGAGATTCAGTGTTGAGATAGCAAAATTCTGCCCCTCGTTCCTTGGCAGGGCCCTATGATT
TATGCAGGAGCAGAGGCAGCACGCAATCGAGCT**GTCA**AGAGAGC**GTCA**GCTATTAGGCAAATGCTGCGTGG
TTTTTGAAGAGGTCGACACTATAAAATCCCACTCCAGGCTCTGGAGTGG <- **RFR 1**

B.

Binding Site	# of Sites	# of sites preserved in ovine
GCCAAT	1	none
CCAAT	2	1
CACCC	4	1
TGAC/GTCA	9	7
TGACGTCAC	1	1

Fig. 4. Regulatory signals in the 5′ RFR 1 and RFR 2 of human CRH gene. (A) The positions of putative consensus binding sites found for transcription factors are underlined and those preserved in the ovine gene are typed in bold face letters. The broken superscript line marks the nonpreserved sequences in the ovine GCCAAT site. (B) List of the type and number of binding sites found in the human CRH gene and their respective ovine counterparts. The sites shown are for the transcription factors CTF/NP1, CP1, CACCC, c-jun/AP-1 and CREB, respectively.

Acknowledgements: We are grateful to Drs Arlene Wyman and David Botstein for their generous gift of the human genomic DNA library. N.C.V. is a research scholar of the Greek Ministry of Research and Technology.

REFERENCES

[1] Chrousos, G.P., Loriaux, D.L. and Gold, P.W. (1988) in: Advances in Experimental Medicine and Biology, vol. 245, Plenum Press, New York.
[2] Gold, P.W., Goodwin, F. and Chrousos, G.P. (1988) N. Engl. J. Med. 319, 348–353.
[3] Gold, P.W., Goodwin, F. and Chrousos, G.P. (1988) N. Engl. J. Med. 319, 413–420.
[4] Gold, P.W., Loriaux, D.L., Roy, A., Kling, M.A., Calabrese, J.R., Kellner, C.H., Nieman, L.K., Post, R.M., Pickar, D., Galluci, W., Avgerinos, P., Paul, S., Oldfield, E.H., Cutler, G.B. Jr. and Chrousos, G.P. (1986) N. Engl. J. Med. 314, 1329–1335.
[5] Gold, P.W., Gwirtzman, H., Avgerinos, P., Nieman, L.K., Gallucci, W.T., Kaye, W., Jimerson, D., Ebert, M., Rittmaster, R., Loriaux, D.L. and Chrousos, G.P. (1986) N. Engl. J. Med. 314, 1335–1342.
[6] Sternberg, E., Hill, J.M., Chrousos, G.P., Kamilaris, T., Listwak, S., Gold, P.W. and Wilder, R.L. (1989) Proc. Natl. Acad. Sci. USA 86, 2374–2378.
[7] Sternberg, E.M., Young, W.S. Jr, Bernardini, R., Calogero, A., Chrousos, G.P., Gold, P.W. and Wilder, R.L. (1989) Proc. Natl. Acad. Sci. USA 86, 4771–4775.
[8] Vale, W., Spiess, J., Rivier, C. and Rivier, J. (1981) Science 213, 1394–1397.
[9] Grino, M., Chrousos, G.P. and Margioris, A.N. (1987) Biochem. Biophys. Res. Commun. 148, 1208–1214.
[10] Frim, D.M., Emanuel, R.L., Robinson, B.G., Smas, C.M., Adler, G.K. and Majzoub, J.A. (1988) J. Clin. Invest. 82, 287–292.
[11] Robinson, B.G., Emanuel, R.L., Frim, D.M. and Majzoub, J.A. (1988) Proc. Natl. Acad. Sci. USA 85, 5244–5248.
[12] Calogero, A.E., Galluci, W.T., Gold, P.N. and Chrousos, G.P. (1988) J. Clin. Invest. 82, 767–774.
[13] Thompson, R.C., Seasholtz, A.F. and Herbert, E. (1987) Mol. Endocrinol. 1, 363–370.
[14] Roche, R.J., Crawford, R.J., Fernley, R.T., Tregear, G.W. and Coghlan, J.P. (1988) Gene 71, 421–431.
[15] Rimm, D.L., Horness, D., Kucera, J. and Blattner, F.R. (1980) Gene 12, 301–309.
[16] Wyman, A.R., Wolfe, L.B. and Botstein, D. (1985) Proc. Natl. Acad. Sci. USA 82, 2880–2884.
[17] Wyman, A.R., Wertman, K.F., Barker, D., Helny, C. and Petri, W.H. (1986) Gene 49, 263–271.
[18] Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning, a Laboratory Manual, 2nd ed, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.
[19] Shibahara, S., Morimoto, Y., Furutani, Y., Notake, M., Takahashi, H., Shimizu, S. and Numa, S. (1983) EMBO J. 2, 775–779.
[20] Chen, E.J. and Seeburg, P.H. (1985) DNA 4, 165–170.
[21] Benton, W.D. and Davis, R.W. (1977) Science 196, 180–182.
[22] Adler, G.K., Smas, C.M. and Majzoub, J.A. (1988) J. Biol. Chem. 263, 5846–5852.
[23] Dorin, R.I., Takahashi, H., Nakai, Y., Fukata, J., Naitoh, Y. and Imura, H. (1989) Mol. Endocrinol. 3, 1537–1544.
[24] Seasholtz, A.F., Thompson, R.C. and Douglas, J.O. (1988) Mol. Endocrinol. 2, 1311–1319.
[25] Thompson, R.C., Seasholtz, A.F., Douglas, J.O. and Herbert, E. (1987) Ann. NY Acad. Sci. 512, 1–11.
[26] Takahashi, H., Hakamata, Y., Watanabe, Y., Kikuno, R., Miyata, T. and Numa, S. (1983) Nucleic Acids Res. 11, 6847–6858.
[27] Barta, A., Richards, R.I., Baxter, J.D. and Shine, J. (1981) Proc. Natl. Acad. Sci. USA 78, 4867–4871.
[28] Yamamoto, K.R., Stallkup, M.R., Ring, J. and Ringold, G.M. (1978) Cold Spring Harbor Symp. Quant Biol. 42, 625–638.
[29] Jingami, H., Matsukura, S., Numa, S. and Imura, H. (1985) Endocrinology 117, 1314–1320.
[30] Young, W.S., Mezey, E. and Sigel, R.E. (1986) Neurosci. Lett. 70, 198–203.
[31] Jantzen, H.M., Strahle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R. and Schutz, G. (1987) Cell 49, 29–38.
[32] Hollenberg, S.M. and Evans, R.M. (1988) Cell 55, 899–906.
[33] Mitchell, P.J. and Tjian, R. (1989) Science 245, 371–378.
[34] Shule, R., Muller, M., Murakami, H.O. and Renkawitz, R. (1988) Nature 323, 87–90.
[35] Ingraham, H.A., Chen, R., Mangalam, H.J., Elsholtz, H.P., Flynn, S.E., Lin, C.R., Simmons, D.M., Swanson, L. and Rosenfeld, G. (1988) Cell 55, 519–529.
[36] Luger, A., Calogero, A., Kalogeras, K., Gallucci, W., Gold, P.W., Loriaux, D.L. and Chrousos, G.P. (1987) J. Clin. Endocrinol. Metab. 66, 334–337.